

in A in  
(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG



(19) Weltorganisation für geistiges Eigentum  
Internationales Büro

(43) Internationales Veröffentlichungsdatum  
20. Juni 2002 (20.06.2002)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 02/49003 A1

(51) Internationale Patentklassifikation<sup>7</sup>: G10L 13/02 (72) Erfinder; und  
(75) Erfinder/Anmelder (nur für US): DAFCIK, Juan  
(21) Internationales Aktenzeichen: PCT/DE01/04658 [DE/DE]; Kornblumenweg 6, 85375 Neufahrn (DE).  
(22) Internationales Anmelde datum: 11. Dezember 2001 (11.12.2001) MENTZ, Stephan [DE/DE]; Uhlandstr. 27, 85386 Eching (DE). VOLLAND, Rainer [DE/DE]; Holzstr. 49, 80469 München (DE).  
(25) Einreichungssprache: Deutsch (74) Gemeinsamer Vertreter: SIEMENS AKTIENGESELLSCHAFT; Postfach 22 16 34, 80506 München (DE).  
(26) Veröffentlichungssprache: Deutsch  
(30) Angaben zur Priorität: 100 62 379.4 14. Dezember 2000 (14.12.2000) DE (81) Bestimmungsstaaten (national): CN, HU, US.  
(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): SIEMENS AKTIENGESELLSCHAFT [DE/DE];  
Wittelsbacherplatz 2, 80333 München (DE). (84) Bestimmungsstaaten (regional): europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

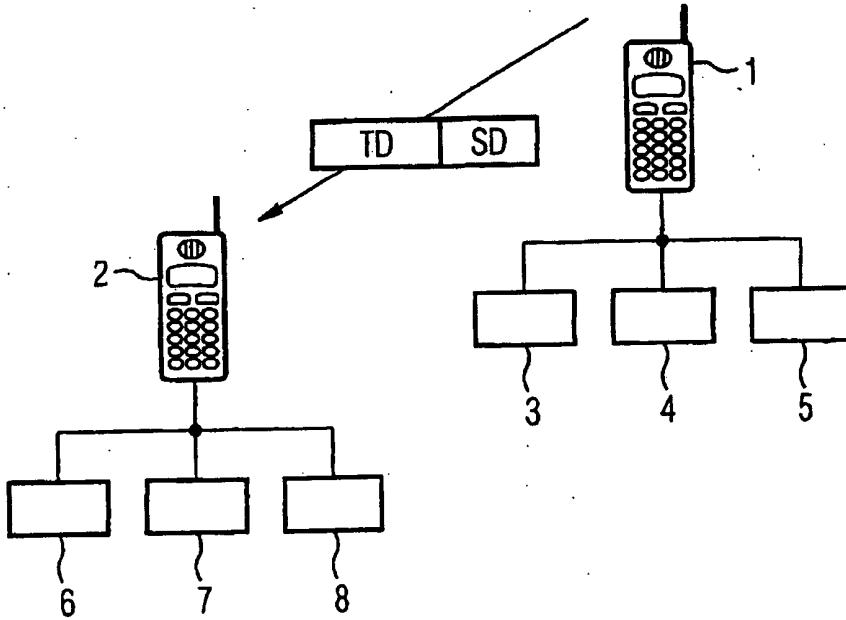
[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD AND SYSTEM FOR CONVERTING TEXT TO SPEECH

(54) Bezeichnung: VERFAHREN UND SYSTEM ZUM UMSETZEN VON TEXT IN SPRACHE



WO 02/49003 A1



(57) Abstract: The aim of the invention is to have texts that are present in electronic form individually read out or replayed by way of the so-called "text-to-speech" technology. To this end, voice parameters (SD) are transmitted together with the text (TD) present in electronic form to a receiving unit (2) provided for replaying the text (TD). Said receiving unit (2) converts the received text (TD) to speech using the voice parameters (SD) received together with the text. The invention makes it possible to have texts, for example SMS messages or e-mails, which are present in electronic form read out with the voice of the sender.

[Fortsetzung auf der nächsten Seite]

**Erklärungen gemäß Regel 4.17:**

- *hinsichtlich der Berechtigung des Anmelders, ein Patent zu beantragen und zu erhalten (Regel 4.17 Ziffer ii) für die folgenden Bestimmungsstaaten CN, HU, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR)*
- *Erfindererklärung (Regel 4.17 Ziffer iv) nur für US*

**Veröffentlicht:**

- *mit internationalem Recherchenbericht*

— *vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen*

*Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.*

---

**(57) Zusammenfassung:** Zum persönlichen Vorlesen bzw. Wiedergeben von in elektronischer Form vorliegenden Texten mittels der sogenannten "Text-to-Speech"-Technologie wird erfahrungsgemäß vorgeschlagen, zusammen mit dem in elektronischer Form vorliegenden Text (TD) Stimmenparameter (SD) an eine zur Wiedergabe des Texts (TD) vorgesehene Empfangseinheit (2) zu übertragen, wobei die Empfangseinheit (2) den empfangenen Text (TD) unter Verwendung der zusammen mit dem Text empfangenen Stimmenparameter (SD) in Sprache umsetzt. Auf diese Weise ist das Vorlesen von in elektronischer Form vorliegenden Texten, wie beispielsweise von SMS-Mitteilungen oder E-Mails, mit der Stimme des Absenders möglich.

## Beschreibung

## Verfahren und System zum Umsetzen von Text in Sprache

5 Die vorliegende Erfindung betrifft ein Verfahren zum Umsetzen von in elektronischer Form vorliegendem Text in Sprache nach dem Oberbegriff des Anspruches 1 sowie ein entsprechend ausgestaltetes System nach dem Oberbegriff des Anspruches 9.

10 Werden auf einem elektronischen Gerät Texte in elektronischer Form gespeichert, so ist es in vielen Fällen günstig, diese von dem entsprechenden Gerät vorlesen zu lassen. Dies trifft insbesondere auf Situationen zu, bei denen man den Text selber nicht ablesen kann, wie beispielsweise beim Autofahren, 15 bei Gebrauch eines Kopfhörers oder aber auch bei sehbehinderten Menschen. Das Vorlesen des in elektronischer Form vorliegenden Textes erfolgt durch die sogenannte "Text-to-Speech"-Technologie (TTS), bei der die Textdaten durch synthetische Spracherzeugung in Sprache umgesetzt werden.

20 Bei herkömmlichen elektronischen Geräten ist die Stimme, mit welcher der in elektronischer Form vorliegende Text vorgelesen wird, sehr unpersönlich und in der Regel für jeden Text gleich. Besonders bei persönlichen Texten, wie beispielsweise 25 bei SMS-Mitteilungen ("Short Message Service") oder bei elektronischer Post ("E-Mail"), wäre ein Vorlesen beispielsweise mit der Stimme des Verfassers oder Absenders des jeweiligen Textes wünschenswert, um somit dem vorgelesenen Text eine persönliche Note zu verleihen.

30 Zur Lösung dieses Problems sind bisher keine Ansätze bekannt. Bei herkömmlichen elektronischen Geräten mit "Text-to-Speech"-Technologie, wie beispielsweise Mobiltelefonen, welche zur Sprachwiedergabe von SMS-Mitteilungen oder E-Mails 35 ausgestaltet sind, besteht lediglich die Möglichkeit, die zur Wiedergabe des in elektronischer Form vorliegenden Textes vorgesehene Stimme grob einzustellen. So kann beispielsweise

zwischen einer männlichen und einer weiblichen Stimme oder zwischen einer Stimme einer jüngeren Person und einer Stimme einer älteren Person ausgewählt werden.

- 5 Der vorliegenden Erfindung liegt daher die Aufgabe zugrunde, eine Möglichkeit bereitzustellen, um in elektronischer Form vorliegende Texte mit einer gewünschten Stimme einer individuellen Person wiederzugeben.
- 10 Diese Aufgabe wird erfindungsgemäß durch ein Verfahren mit den Merkmalen des Anspruches 1 bzw. ein System mit den Merkmalen des Anspruches 9 gelöst. Die Unteransprüche definieren jeweils bevorzugte und vorteilhafte Ausführungsformen der vorliegenden Erfindung.
- 15 Erfindungsgemäß wird vorgeschlagen, dem jeweils vorzulesenden Text einen Stimmenparametersatz zuzuordnen, welcher eine für die Wiedergabe des Textes vorgesehene Stimme einer zuvor ausgewählten Person beschreibt. Dabei kann es sich insbesondere um die Stimme des Erstellers bzw. Absenders des jeweiligen, in elektronischer Form vorliegenden Textes handeln. Auf diese Weise könnte beispielsweise eine empfangene SMS-Mitteilung von einem Mobiltelefon mit der Stimme des Absenders vorgelesen werden. Hierzu ist lediglich erforderlich, dass auf dem 20 Sendegerät, an dem der Text erstellt wird, einmalig eine Extraktion der Stimmenparameter des Texterstellers bzw. Absenders vorgenommen wird, wobei die Stimmenparameter dann für jeden von derselben Person erstellten Text auf Wunsch zusammen mit den Textdaten an das jeweilige Empfangsgerät gesendet 25 werden kann.
- 30

Gemäß einer Variante der vorliegenden Erfindung ist auf dem Empfangsgerät bereits eine Datenbank mit Stimmenparametern verschiedener Kommunikationspartner gespeichert. In diesem Fall ist lediglich erforderlich, von dem Sendegerät zusammen mit den Textdaten eine Kennung ("Identifier", ID) zu übertragen, welche den jeweiligen Absender oder Ersteller des Textes

bzw. deren Stimme bezeichnet. Im Empfangsgerät wird dann an-  
hand der zusammen mit den Textdaten empfangenen Kennung über-  
prüft, ob bereits ein Stimmenparameter-Datensatz für die  
durch die Kennung bezeichnete Person vorliegt. Ist dies der  
5 Fall, werden die entsprechenden Stimmenparameter ausgelesen  
und für die Wiedergabe des in elektronischer Form vorliegen-  
den Textes verwendet. Ist dies jedoch nicht der Fall, kann an  
das Sendegerät ein Befehl zurückgesendet werden, wodurch vom  
Sendegerät die gewünschten Stimmenparameter angefordert wer-  
10 den.

Im Rahmen der vorliegenden Erfindung ist nicht nur vorgese-  
hen, für die Wiedergabe des in elektronischer Form vorliegen-  
den Textes die Stimmenparameter des Texterstellers zu verwen-  
den. Vielmehr können vom Sendegerät auch Stimmenparameter be-  
liebiger Personen zusammen mit den Textdaten übertragen wer-  
den, um eine bestimmte Wirkung zu erzielen. Dies ist bei-  
spielsweise insbesondere dann vorteilhaft, wenn der Texter-  
steller nicht mit der eigentlichen Person des Absenders über-  
20 einstimmst (beispielsweise für den Fall, dass eine Sekretärin  
eine E-Mail für ihren Chef verfasst und im Namen des Chefs  
versendet). Ebenso können zur Wiedergabe des Textes Stimmen-  
parameter bekannter, prominenter Personen (beispielsweise von  
Schauspielern, Politikern etc.) vorgegeben werden.

25 Die vorliegende Erfindung eignet sich insbesondere für die  
Versendung und sprachliche Wiedergabe von E-Mails, SMS-  
Mitteilungen oder elektronischen Grußkarten über Computer o-  
der mobile Endgeräte, wie beispielsweise Mobiltelefone,  
30 Schnurlosetelefone oder Handheld-Organizer ("PDA"). Ebenso  
kann die vorliegende Erfindung auch zur Realisierung einer  
persönlichen Vorlesestimme zum Vorlesen von elektronischen  
Büchern verwendet werden. Die vorliegende Erfindung ist je-  
doch nicht auf den Anwendungsfall des Versendens von in e-  
35 lektronischer Form vorliegenden Texten (beispielsweise über  
kabellose Funkverbindungen) beschränkt, sondern kann grund-  
sätzlich auch auf Anwendungsfälle angewendet werden, bei de-

nen der vorzulesende elektronische Text bereits in dem jeweiligen Wiedergabegerät gespeichert ist.

Die vorliegende Erfindung wird nachfolgend anhand bevorzugter 5 Ausführungsbeispiele unter Bezugnahme auf die beigefügte Zeichnung näher erläutert.

Figur 1 zeigt eine schematische Darstellung zur Erläuterung 10 der Übertragung und sprachlichen Wiedergabe von Textdaten gemäß einem ersten Ausführungsbeispiel der vorliegenden Erfindung, und

Figur 2 zeigt eine schematische Darstellung zur Erläuterung 15 der Übertragung und sprachlichen Wiedergabe von Textdaten gemäß einem zweiten Ausführungsbeispiel der vorliegenden Erfindung.

In Figur 1 ist eine Kommunikation zwischen zwei Mobiltelefonen 1, 2 dargestellt. Dabei wird davon ausgegangen, dass an 20 dem Mobiltelefon 1 eine auch als SMS-Mitteilung bezeichnete Kurzmitteilung erstellt und über Funkverbindung an das Mobiltelefon 2 übertragen wird. Von dem Mobiltelefon 2 wird der in elektronischer Form vorliegende Text der SMS-Mitteilung mittels der sogenannten "Text-to-Speech"-Technologie durch synthetische Spracherzeugung in Sprache umgesetzt und über den 25 Lautsprecher des Mobiltelefons 2 wiedergegeben.

Bei dem in Figur 1 gezeigten Ausführungsbeispiel umfasst das Mobiltelefon 1 eine Einheit 3, welche eine Sprachprobe bestimmter Länge des Benutzers des Mobiltelefons 1 aufnimmt und 30 daraus bestimmte Stimmenparameter des Benutzers des Mobiltelefons 1 extrahiert, wobei die auf diese Weise gewonnenen Stimmenparameter möglichst exakt die Stimme des Benutzers des Mobiltelefons 1 beschreiben bzw. charakterisieren. Die auf 35 diese Weise gewonnenen Stimmenparameter werden in einem Stimmenparameter-Speicher 4 gespeichert, so dass im Prinzip lediglich eine einmalige Extraktion der Stimmenparameter des

Benutzers des Mobiltelefons 1 erforderlich ist. Wird nunmehr von dem Benutzer des Mobiltelefons 1 eine SMS-Mitteilung durch Tasteneingabe erstellt, wird diese in einem SMS-Speicher 5 zwischengespeichert und anschließend auf Tastendruck in Form entsprechender Textdaten TD an das Mobiltelefon 2 des gewünschten Empfängers gesendet. Dabei ist das Mobiltelefon 1 derart ausgestaltet, dass auf Wunsch nicht nur die Textdaten TD der SMS-Mitteilung übertragen werden, sondern auch die dem Ersteller der SMS-Mitteilung zugeordneten und aus dem Stimmenparameter-Speicher 4 ausgelesenen Stimmenparameter. Diese Stimmenparameter können beispielsweise wie in Figur 1 gezeigt in Form entsprechender Stimmenparameterdaten SD an die Textdaten TD der SMS-Mitteilung angehängt und an das Mobiltelefon 2 übertragen werden.

Das empfangende Mobiltelefon 2 ist derart ausgestaltet, dass bei Empfang einer SMS-Mitteilung die Textdaten TD der empfangenen SMS-Mitteilung zunächst in einem SMS-Speicher 8 zwischengespeichert werden. Darüber hinaus ist eine TTS-Einheit 6 vorgesehen, welche auf Tastendruck mittels synthetischer Spracherzeugung unter Anwendung der sogenannten "Text-to-Speech"-Technologie (TTS) eine empfangene und gespeicherte SMS-Mitteilung in Sprache umsetzt und über den Lautsprecher des Mobiltelefons 2 wiedergibt. Dabei erfolgt die Wiedergabe der SMS-Mitteilung in Abhängigkeit von den zusammen mit den entsprechenden Textdaten übertragenen Stimmenparameterdaten SD, welche bei Empfang der entsprechenden SMS-Mitteilung in einen Stimmenparameter-Speicher 7 zwischengespeichert und für eine spätere Sprachwiedergabe von der TTS-Einheit wieder ausgelesen werden können. Der in elektronischer Form vorliegende Text der empfangenen SMS-Mitteilung wird somit derart von der TTS-Einheit 6 mittels synthetischer Spracherzeugung unter Berücksichtigung der entsprechenden Stimmenparameterdaten in Sprache umgesetzt, dass der Text der SMS-Mitteilung mit der Stimme des Verfassers bzw. Absenders der SMS-Mitteilung oder einer möglichst ähnlichen Stimme wiedergegeben wird.

Das in Figur 2 dargestellte Ausführungsbeispiel ist in vielen Punkten ähnlich zu dem in Figur 1 dargestellten Ausführungsbeispiel. Es ist wieder eine Sendeeinheit in Form eines Mobiltelefons 1 vorgesehen, welche eine mit dem Mikrofon des

5 Mobiltelefons 1 gekoppelte Einheit 3 zur Extraktion der Stimmenparameter des jeweiligen Benutzers des Mobiltelefons 1, einen Stimmenparameter-Speicher 4 zum Speichern der somit gewonnenen Stimmenparameter und einen SMS-Speicher 5 zum Speichern von an dem Mobiltelefon 1 erstellten SMS-Mitteilungen

10 umfasst. Ebenso ist eine Empfangseinheit in Form eines Mobiltelefons 2 vorgesehen, welche eine TTS-Einheit 6 zur Umsetzung von in elektronischer Form vorliegenden Text mittels synthetischer Spracherzeugung in Sprache und einen SMS-Speicher 8 zum Speichern von empfangenen SMS-Mitteilungen

15 bzw. der entsprechenden Textdaten TD umfasst. Darüber hinaus weist das Mobiltelefon 2 anstelle des in Figur 1 gezeigten Stimmenparameter-Speichers 7 einen Stimmenparameter-Speicher 9 auf, in dem bereits eine Datenbank mit Stimmenparametern verschiedener Kommunikationspartner gespeichert ist.

20 Bei dem in Figur 2 gezeigten Ausführungsbeispiel werden nach Erstellung einer SMS-Mitteilung an dem Mobiltelefon 1 die Textdaten TD der SMS-Mitteilung wie bei dem in Figur 1 gezeigten Ausführungsbeispiel an das Mobiltelefon 2 übertragen.

25 Im Gegensatz zu dem in Figur 1 gezeigten Ausführungsbeispiel werden jedoch bei dem in Figur 2 gezeigten Ausführungsbeispiel nicht die Stimmenparameter des Verfassers der SMS-Mitteilung an das Mobiltelefon 2 übertragen, sondern es wird zusammen mit den Textdaten TD der SMS-Mitteilung eine Kennung

30 ID übertragen, welche das Mobiltelefon 1 bzw. den Benutzer oder Verfasser der SMS-Mitteilung eindeutig charakterisiert. Dabei kann es sich beispielsweise im einfachsten Fall um die Rufnummer des Mobiltelefons 1 handeln, welche ohnehin in digitalen Kommunikationssystemen an das jeweils gerufene Endgerät mit übertragen wird.

35

Bei Empfang einer SMS-Mitteilung von dem Mobiltelefon 1 werden die entsprechenden Textdaten TD in dem Mobiltelefon 2 bzw. in dem SMS-Speicher 8 zwischengespeichert. Darüber hinaus wird zu der empfangenen SMS-Mitteilung die Kennung ID des jeweiligen Absenders bzw. Verfassers der SMS-Mitteilung abgespeichert.

Wird von dem Benutzer des Mobiltelefons 2 eine Sprachwiedergabe der empfangenen und zwischengespeicherten SMS-Mitteilung gewünscht, überprüft die TTS-Einheit 6, ob zu der entsprechenden Kennung ID bereits ein Stimmenparametersatz in dem Stimmenparameter-Speicher 9 gespeichert ist. Ist dies der Fall, werden die der Kennung ID entsprechenden Stimmenparameter aus dem Stimmenparameter-Speicher 9 ausgelesen und von der TTS-Einheit der Spracherzeugung bzw. Sprachwiedergabe der Textdaten TD der jeweiligen SMS-Mitteilung zugrundegelegt. Ist hingegen für die jeweilige Kennung ID, d.h. für die der jeweiligen Kennung ID entsprechende Person, noch kein Stimmenparametersatz in dem Stimmenparameter-Speicher 9 gespeichert, veranlasst die TTS-Einheit 6, dass von dem Mobiltelefon 2 eine Aufforderung an dasjenige Mobiltelefon 1 gesendet wird, von dem die wiederzugebende SMS-Mitteilung stammt, um somit von dem Mobiltelefon 1 die fehlenden Stimmenparameter anzufordern. Bei Empfang einer derartigen Aufforderung muss dann, wie zuvor anhand von Figur 1 beschrieben worden ist, mit Hilfe der Einheit 3 eine Stimmenprobe des Benutzers des Mobiltelefons 1 aufgenommen und daraus die entsprechenden Stimmenparameter extrahiert werden. Anschließend können die Stimmenparameter über die Funkverbindung an das Mobiltelefon 2 gesendet werden, wo sie anschließend in dem Stimmenparameter-Speicher 9 in Form eines entsprechenden Stimmenparametersatzes gespeichert werden und somit für die gewünschte Sprachwiedergabe der bereits zuvor von dem Mobiltelefon 1 an das Mobiltelefon 2 übertragenen SMS-Mitteilung durch die TTS-Einheit 6 zur Verfügung stehen.

## Patentansprüche

1. Verfahren zum Umsetzen von Text in Sprache,  
wobei von einer Sendeeinheit (1) ein in elektronischer Form  
5 vorliegender Text (TD) an eine Empfangseinheit (2) übertragen  
und von der Empfangseinheit (2) mittels synthetischer Sprach-  
erzeugung in Sprache umgesetzt und wiedergegeben wird,  
dadurch gekennzeichnet,  
dass von der Sendeeinheit (1) zusammen mit dem in elektroni-  
10 scher Form vorliegenden Text (TD) personengebundene Informa-  
tionen (SD, ID) an die Empfangseinheit (2) übertragen werden,  
und  
dass von der Empfangseinheit (2) aus den personengebundenen  
Informationen (SD, ID) eine für die Wiedergabe des Textes  
15 (TD) vorgesehene Stimme abgeleitet und bei der Umsetzung so-  
wie Wiedergabe des Textes (TD) verwendet wird.

2. Verfahren nach Anspruch 1,  
dadurch gekennzeichnet,  
20 dass von der Sendeeinheit (1) als personengebundene Informa-  
tionen personengebundene Stimmenparameter (SD) zusammen mit  
dem in elektronischer Form vorliegenden Text (TD) an die Emp-  
fangseinheit (2) übertragen werden, wobei die personengebun-  
denen Stimmenparameter (SD) eine für die Wiedergabe des Tex-  
25 tes (TD) vorgesehene Stimme einer bestimmten Person beschrei-  
ben, und  
dass von der Empfangseinheit (2) der Text (TD) in Abhängig-  
keit von den personengebundenen Stimmenparametern (SD) in  
Sprache umgesetzt und mit der entsprechenden Stimme wiederge-  
30 geben wird.

3. Verfahren nach Anspruch 2,  
dadurch gekennzeichnet,  
dass von der Sendeeinheit (1) die Stimme der den Text (TD)  
35 erstellenden Person in die personengebundenen Stimmenparame-  
ter (SD) zerlegt und in Form der personengebundenen Stimmen-  
parameter (SD) an die Empfangseinheit (2) übertragen wird.

4. Verfahren nach Anspruch 1,  
dadurch gekennzeichnet,  
dass in der Empfangseinheit (2) mehrere Datensätze von perso-  
nengebundenen Stimmenparametern (SD), welche die Stimmen un-  
terschiedlicher Personen beschreiben, gespeichert werden,  
dass von der Sendeeinheit (1) zusammen mit dem in elektroni-  
scher Form vorliegenden Text (TD) als personen gebundene In-  
formationen eine für die Wiedergabe des Textes bestimmte Per-  
son oder deren Stimme beschreibende Kennung (ID) an die Emp-  
fangseinheit (2) übertragen wird, und  
dass in der Empfangseinheit (2) die der Kennung (ID) entspre-  
chenden Stimmenparameter ausgelesen und für die Umsetzung des  
in elektronischer Form vorliegenden Textes (TD) in Sprache  
verwendet werden.

5. Verfahren nach Anspruch 4,  
dadurch gekennzeichnet,  
dass von der Empfangseinheit (2) von der Sendeeinheit (1) die  
20 der zuvor von der Sendeeinheit (1) übertragenen Kennung (ID)  
entsprechenden Stimmenparameter angefordert werden, falls in  
der Empfangseinheit (2) noch kein der Kennung (ID) entspre-  
chender Datensatz von personen gebundenen Stimmenparametern  
gespeichert ist.

25 6. Verfahren nach Anspruch 4 oder 5,  
dadurch gekennzeichnet,  
dass die von der Sendeeinheit (1) an die Empfangseinheit (2)  
zusammen mit dem in elektronischer Form vorliegenden Text  
30 (TD) übertragene Kennung (ID) diejenige Person oder deren  
Stimme bezeichnet, welche den Text (TD) auf der Sendeeinheit  
(1) erstellt hat.

7. Verfahren nach einem der vorhergehenden Ansprüche,  
35 dadurch gekennzeichnet,  
dass als Empfangseinheit (2) ein mit der Sendeeinheit (1) ka-  
bellos kommunizierendes mobiles Gerät (2) verwendet wird.

8. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass als Sendeeinheit (1) ein mit der Empfangseinheit (2) kabellos kommunizierendes mobiles Gerät verwendet wird.

5

9. System zum Umsetzen von Text in Sprache, mit einer Sendeeinheit (1) zum Erstellen eines in elektronischer Form vorliegenden Texts (TD) und zur Übertragung des in elektronischer Form vorliegenden Texts (TD) an eine Empfangseinheit (2),

10 wobei die Empfangseinheit (2) Spracherzeugungsmittel (6) zur Umsetzung des von der Sendeeinheit (1) empfangenen und in elektronischen Form vorliegenden Texts (TD) in Sprache umfasst,

15 dadurch gekennzeichnet, dass die Sendeeinheit (1) zur Übertragung von personengebundenen Informationen (SD, ID) zusammen mit dem in elektronischer Form vorliegenden Text (TD) an die Empfangseinheit (2) ausgestaltet ist, und

20 dass die Spracherzeugungsmittel (6) der Empfangseinheit (2) derart ausgestaltet sind, dass sie aus den von der Sendeeinheit (1) empfangenen personengebundenen Informationen (SD, ID) personengebundene Stimmenparameter, welche eine für die

25 Wiedergabe des Texts (TD) vorgesehene Stimme einer bestimmten Person beschreiben, ableiten und die Umsetzung des Texts in die Sprache auf Grundlage dieser personengebundenen Stimmenparameter durchführen.

30 10. System nach Anspruch 9, dadurch gekennzeichnet, dass die Sendeeinheit (1) zur Übertragung des in elektronischer Form vorliegenden Texts (TD) zusammen mit den personengebundenen Stimmenparametern (SD) als personengebundene Informationen an die Empfangseinheit (2) ausgestaltet ist, und

35 dass die Spracherzeugungsmittel (6) der Empfangseinheit (2) derart ausgestaltet sind, dass sie den von der Empfangseinheit

heit (1) empfangenen Text (TD) in Abhängigkeit von den ebenfalls von der Empfangseinheit (1) empfangenen personengebundenen Stimmenparametern (SD) in Sprache umsetzen und mit der entsprechenden Stimme wiedergeben.

5

11. System nach Anspruch 10,  
dadurch gekennzeichnet,  
dass die Sendeeinheit (1) Stimmenparametererzeugungsmittel  
(3) zur Erzeugung von Stimmenparametern (SD), welche die  
10 Stimme der den Text (TD) erstellenden Person beschreiben, umfassen.

12. System nach Anspruch 9,  
dadurch gekennzeichnet,  
15 dass die Empfangseinheit (2) Speichermittel (9) zum Speichern  
mehrerer Datensätze von personengebundenen Stimmenparametern,  
welche die Stimmen unterschiedlicher Personen beschreiben,  
umfassen, und  
dass die Sendeeinheit (1) derart ausgestaltet ist, dass sie  
20 zusammen mit dem in elektronischer Form vorliegenden Text  
(TD) als personengebundene Information eine Kennung (ID) an  
die Sendeeinheit (2) überträgt, wobei diese Kennung (ID) die  
jenige Person, deren Stimme für die Wiedergabe des Textes  
(TD) vorgesehen ist, oder deren Stimme beschreibt und  
25 dass die Spracherzeugungsmittel (6) der Empfangseinheit (2)  
derart ausgestaltet sind, dass sie den der von der Sendeeinheit  
(1) empfangenen Kennung entsprechenden Datensatz von  
personengebundenen Stimmenparametern aus den Speichermitteln  
(9) auslesen und für die Umsetzung des von der Sendeeinheit  
30 (1) empfangenen Texts (TD) in Sprache zugrundelegen.

13. System nach Anspruch 12,  
dadurch gekennzeichnet,  
dass die Empfangseinheit (2) derart ausgestaltet ist, dass  
35 sie für den Fall, dass in den Speichermitteln (9) kein der  
von der Sendeeinheit (1) empfangenen Kennung entsprechender  
Datensatz von personengebundenen Stimmenparametern gespei-

12

chert ist, einen Befehl zur Anforderung von der Kennung (ID) entsprechenden personengebundenen Stimmenparametern an die Sendeeinheit (1) sendet.

5 14. System nach Anspruch 12 oder 13,  
dadurch gekennzeichnet,  
dass die von der Sendeeinheit (1) an die Empfangseinheit (2)  
übertragene Kennung (ID) diejenige Person bezeichnet, welche  
zuvor an der Sendeeinheit (1) den Text (TD) erstellt hat, o-  
10 der deren Stimme bezeichnet.

1/1

FIG 1

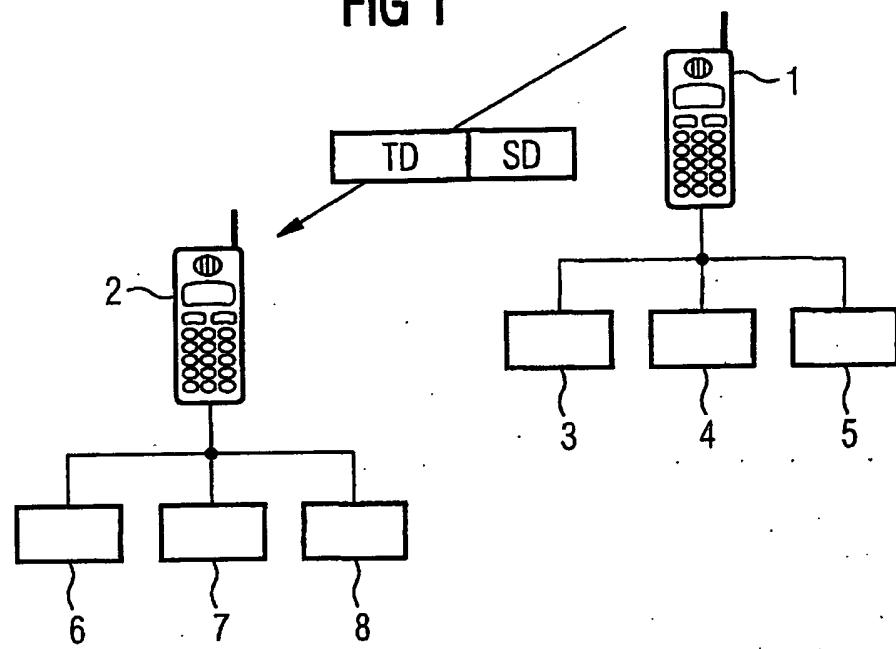


FIG 2

